A polyhedral Frobenius theorem with applications to integer optimization

David Adjiashvili
Institute for Operations Research, ETH Zürich
Rämistrasse 101, 8092 Zürich, Switzerland
email: david.adjiashvili@ifor.math.ethz.ch

Timm Oertel
Institute for Operations Research, ETH Zürich
Rämistrasse 101, 8092 Zürich, Switzerland
email: timm.oertel@ifor.math.ethz.ch

Robert Weismantel
Institute for Operations Research, ETH Zürich
Rämistrasse 101, 8092 Zürich, Switzerland
email: robert.weismantel@ifor.math.ethz.ch

Abstract

We prove a representation theorem of projections of sets of integer points by an integer matrix W. Our result can be seen as a polyhedral analogue of several classical and recent results related to the *Frobenius problem*.

Our result is motivated by a large class of non-linear integer optimization problems in variable dimension. Concretely, we aim to optimize $f(Wx)$ over a set $\mathcal{F} = P \cap \mathbb{Z}^n$, where f is a non-linear function, $P \subset \mathbb{R}^n$ is a polyhedron and $W \in \mathbb{Z}^{d \times n}$. As a consequence of our representation theorem, we obtain a general efficient transformation from the latter class of problems to integer linear programming. Our bounds depends polynomially on various important parameters of the input data leading, among others, to first polynomial time algorithms for several classes of non-linear optimization problems.

1 Introduction

Non-linear integer programming is concerned with optimizing a non-linear function over the integer points in a polyhedron. Significant effort has been made in recent years to extend the well-established theory of *linear* integer programming to the non-linear case. Along these lines, polynomial time algorithms for various classes of nonlinear objective functions were developed, including convex functions [10], bounded-degree polynomials [7, 6] and more. Apart from very few exceptions [12, 5], however, all results in this vein were proved for the *fixed-dimension* case, namely for the case where the total number of variables is a fixed constant. The latter fact makes these methods less practical, limiting their potential domain of applications.

There are, of course, good reasons why positive algorithmic results in non-linear variable-dimension integer programming are harder to come by. Firstly, this class of problems trivially generalizes linear integer programming, which is NP-hard in almost every variable-dimension setup. Secondly, non-linear variable-dimension integer problems often become hard already in the fixed-dimensional case. Finally, if the non-linear function acts directly on the variable-dimensional space, even stronger hardness results can be proved. For example, in the function oracle model one can prove simple information-theoretic exponential lower bounds on the complexity of any algorithm approximating the minimum of a convex function over the hypercube. If the function class is further restricted to be convex quadratic polynomials and stronger oracles are assumed, the latter problem becomes "merely" NP-hard to solve exactly. Worse still, the latter example shows the large increase in complexity when a linear objective function is replaced with a non-linear one. This means that any algorithm reducing the latter problems to integer linear programming will most likely need to replace the well-structured feasible set, namely the hypercube, with a much more complicated one.

Still, it is meaningful to ask: *Which non-linear variable-dimension integer programming problems can be reduced to the linear case, maintaining the structure of the problem class?* In the present paper we study one such class of problems. Our class of problems contains an additional component, namely that of a *projection* into a low-dimensional space. The previous discussion suggests that this is, to a large extent, unavoidable when efficient reductions of the latter type are sought. Formally, we are interested in studying problems of the form

$$\min\{f(Wx) \mid x \in \mathbb{Z}^n \cap P\}, \tag{1}$$

where $f : \mathbb{R}^d \to \mathbb{R}$ is a function from our function class, $P := \{x \in \mathbb{R}^n \mid Ax \leq b\}$ is a polyhedron in n-dimensional space, (with $A \in \mathbb{Z}^{m\times n}, b \in \mathbb{Z}^m$) and W is a $d \times n$ integer matrix. We discuss minimization here, but our results also hold for maximization problems. The set $\mathcal{F} := \{x \in \mathbb{Z}^n \mid x \in P\}$ is called the *feasible set*, and points $x \in \mathcal{F}$ are called *feasible*. Although not necessary for our main result, we think of n as being large (variable) and of d as being small (fixed). We note that this class of problems includes linear integer programming already for $d = 1$ and f–the identity.

In this paper we give a first general-purpose efficient reduction from the latter class of problems to integer programming. The efficiency of our reduction depends on various input parameters. We elaborate on this exact dependence later. As a result, we obtain the first polynomial algorithms for several classes of variable-dimension non-linear integer problems. For other problem classes, our method provides a polynomial time reduction from the non-linear problem to linear integer programming, maintaining the structure of the feasible set.

We assume black-box access to two oracles, namely a *fiber oracle* and a *d-dimensional non-linear optimization oracle* (or simply, *optimization oracle*), defined as follows. In what follows by "opt" we mean both "min" and "max".

Definition 1 (fiber oracle, optimization oracle)**.**

- *The fiber oracle accepts as input a point $y \in \mathbb{Z}^d$ and either returns a feasible point (a point $x \in \mathcal{F}$) such that $Wx = y$, or asserts that no such point exists.*
- *The optimization oracle accepts descriptions of a polyhedron $R \subset \mathbb{R}^d$ and an affine sub-lattice $\Lambda \subset \mathbb{Z}^d$ of the integer lattice, and returns a point y^* in*

$$\arg\operatorname{opt}\{f(y) \mid y \in \Lambda \cap R\},$$

if one exists, or asserts that the latter set is empty.

We note that both oracles can be implemented in polynomial time for various classes of input parameters. We defer a detailed discussion on this topic to a later stage. With Definition 1 in mind, we can now state our main algorithmic result.

Theorem 2. *Let d be any fixed constant. There is an algorithm that solves the non-linear optimization problem*

$$\operatorname{opt}\{f(Wx) \mid Ax \leq b,\ x \in \mathbb{Z}^n\},$$

with input $A \in \mathbb{Z}^{m\times n}$, $W \in \mathbb{Z}^{d\times n}$, $b \in \mathbb{Z}^m$ and $f : \mathbb{R}^d \to \mathbb{R}$. The amount of work and the number of oracle calls it performs (to the optimization and fiber oracles) is polynomial in n, the maximum sub-determinant of A and the unary encoding length of W.

Let us now draw a road map for the proof of the latter theorem. Theorem 2 follows from a careful analysis of the set

$$\mathcal{R} = W\mathcal{F} := \{y = Wx \mid x \in \mathcal{F}\},$$

namely the projection of the feasible set with respect to the matrix W. Let us first explain why understanding this set can have important algorithmic consequences. Assume, for example, that $\mathcal{R} = Q \cap \mathbb{Z}^d$ holds, where

$$Q := WP = \{Wx \mid x \in P\}.$$

In this case we can solve Problem (1) with two oracle calls as follows. First, use the optimization oracle to obtain $y^* \in \arg\min\{f(y) \mid y \in Q \cap \mathbb{Z}^d\}$. This is possible since $\mathbb{Z}^d$ is clearly a lattice, thus the latter problem has the required form. Then, use the fiber oracle to obtain $x^* \in \mathcal{F}$ with $Wx^* = y^*$. The oracle is

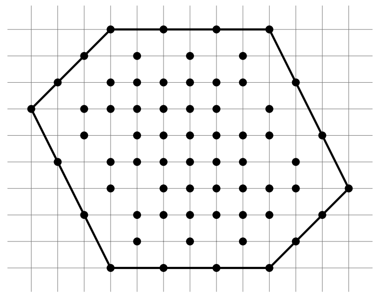

Figure 1: An illustration of the set $\mathcal{R}$ and the notion of holes.

guaranteed to return a point $x^* \in \mathcal{F}$ since we assumed that $\mathcal{R} = Q \cap \mathbb{Z}^d$, namely that *every* integer point in Q has a feasible pre-image under the projection with W. The obtained x^* is clearly an optimal solution. In the following remark we give a concrete example of a class of matrices with this property.

Remark 3. *One important case in which $\mathcal{R} = Q \cap \mathbb{Z}^d$ is the case of a totally unimodular matrix $\binom{W}{A}$ (see e.g. [18, Theorem 19.1]). In this case one can show the inclusion $Q \cap \mathbb{Z}^d \subset \mathcal{R}$ as follows. Let $y \in Q \cap \mathbb{Z}^d$. Since $y \in Q$ there exists $x \in \mathbb{R}^n$ such that $Ax \leq b$ and $Wx = y$. Since $\binom{W}{A}$ is totally unimodular and $y \in \mathbb{Z}^d$, the solution set to the latter system is an integral polyhedron. Thus, there exists an integral point $\bar{x} \in \mathbb{Z}^n$ with $A\bar{x} \leq b$ and $W\bar{x} = y$. This implies that $y \in \mathcal{R}$.*

It is, unfortunately, rarely the case that $\mathcal{R} = Q \cap \mathbb{Z}^d$, as typically one has

$$\mathcal{R} \subsetneq Q \cap \mathbb{Z}^d,$$

namely, the set $Q \cap \mathbb{Z}^d$ contains *holes*, i.e., points without pre-images in $\mathcal{F}$. We illustrate this with the following simple example.

Example 4. *Let $n = 3, d = 2, P = \{x \in \mathbb{R}^3 \mid 0 \leq x_i \leq 3,\ i = 1, 2, 3\}$, and consider the matrix*

$$W = \begin{pmatrix} 1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix}.$$

Figure 1 illustrates the polyhedron Q and the set $\mathcal{R}$, which corresponds to the thick points. All other points are holes.

In this more common situation, the latter simple strategy cannot be directly applied. One can still hope, however, to decompose the problem into subproblems, each solvable in this way. Ideally, this decomposition should have the form

$$\mathcal{R} = \bigcup_{i \in I} Q^i \cap \Lambda^i,$$

where $Q^i \subset \mathbb{R}^d$ are polyhedra and $\Lambda^i \subset \mathbb{Z}^d$ are affine sub-lattices, the description of which can be efficiently computed from the input data. Then, $k = 2|I|$

oracle calls are sufficient to solve the problem, by simply repeating our simple procedure for every sub-problem $i \in I$, defined over $Q^i \cap \Lambda^i$, and taking the best solution, among the $|I|$ resulting candidates. It is hence of particular interest to study such efficient representations, trying to minimize k, while maintaining the property that both Q^i and Λ^i are efficiently computable.

Our main contribution provides such a decomposition. Concretely, we show strong existential bounds on some important parameters of such decompositions. These bounds, in turn, lead to strong bounds on efficiently computable decompositions, which are later exploited to obtain efficient algorithms.

It is now evident that the current paper deals with a problem of representability of sets of integer vectors. Indeed, what we seek in our decomposition is a way to cover all points in $\mathcal{R}$ by "simple" sets, with the property that none of these sets contains a hole, namely a point in $(Q \cap \mathbb{Z}^d) \setminus \mathcal{R}$. One can almost equivalently ask: *How complicated can the set of holes be?*

Our result can hence be seen as a polyhedral variant of the *Frobenius problem*, also known as the *coin problem*. Given a set $S = \{a_1, \dots, a_m\} \subset \mathbb{Z}_+$ of positive integers with $\gcd(S) = 1$, the Frobenius problem asks to find the largest integer $k \in \mathbb{Z}_+$ that can not be represented as a positive integer combination of numbers in S. The Frobenius problem is known to be NP-hard [17] in all but a few special cases [14, 13]. Several bounds on the Frobenius number were also proven [9, 3, 2].

For a positive integer s, let $[s] = \{1, \dots, s\}$. In higher dimensions one can define the following generalization of the Frobenius problem, called the *diagonal Frobenius problem* [1]. Given a $d \times n$ integer matrix M with the property that $\text{cone}(M) = \{M\lambda \mid \lambda \in \mathbb{R}^n_+\}$ forms a full-dimensional pointed cone, and such that $M\mathbb{Z}^n = \mathbb{Z}^d$, find the smallest $t \in \mathbb{Z}$ with the property that

$$(tv + \text{cone}(M)) \cap \mathbb{Z}^d \subset \{Mx \mid x \in \mathbb{Z}^m_+\},$$

where $v = \sum_{i\in[m]} M_{\star,i}$ is the sum of the columns of M. We note that there are several ways to define such a generalization. Intuitively, the diagonal Frobenius number is the smallest factor by which one needs to shift the cone $\text{cone}(M)$ inwards (in the direction $v \in \text{cone}(M)$), so that every integer point in it be expressible as a positive integer combination of the columns of M. The following result of Aliev and Henk [1] proves a strong bound on the diagonal Frobenius number.

Theorem 5 (Aliev and Henk 2010)**.** *Let $M \in \mathbb{Z}^{d\times n}$, such that $M\mathbb{Z}^n = \mathbb{Z}^d$ with* $\text{cone}(M)$ *pointed. Then the diagonal Frobenius number of M is at most*

$$\mathbf{c}(M) = \frac{(n-d)\sqrt{n}}{2}\sqrt{\det(MM^{\mathsf{T}})}.$$

Theorem 5 guarantees that the set $M\mathbb{Z}^n_+$ becomes very regular in the cone $\text{cone}(M)$ shifted by the vector $\mathbf{c}(M)v$. For our purposes we need a similar result for arbitrary polyhedra, instead of cones. To this end we define a notion of regularity, suitable for our needs.

Definition 6 (Δ-regular set)**.** *We call a set $\mathcal{S} \subset \mathbb{Z}^d$ Δ-regular, with respect to a region $B \subset \mathbb{R}^d$, if there exists a family of full-dimensional affine sub-lattices $\Lambda_1, \cdots, \Lambda_k$ of $\mathbb{Z}^d$ with determinants* $\det(\Lambda_i) \leq \Delta$ *such that*

$$\mathcal{S} \cap B = \bigcup_i \Lambda_i \cap B. \tag{2}$$

Theorem 5 can be restated in terms of our new definition as follows. For a matrix M satisfying the conditions of Theorem 5, the set $\mathcal{S} = M\mathbb{Z}_+^n$ is 1-regular with respect to $B = \mathbf{c}(M)v + \text{cone}(M)$. Furthermore, only the lattice $\Lambda = \mathbb{Z}^d$ is needed to certify this fact.

Our main result proves a similar statement for a much more general setup. Firstly, the matrix M satisfying the conditions of Theorem 5 is replaced with the arbitrary matrix W. Secondly, the admissible set of positive combinations is no longer the convenient set $\mathbb{Z}_+^n$, but rather the set $\mathcal{F}$. Finally, we prove regularity with respect to a polyhedron $Q' \subset Q$. Since Q can be bounded, Q' can no longer be a translate of Q. We use instead the notion of *α-inscribed polyhedron* defined as follows. Let $R \subset \mathbb{R}^d$ be a polyhedron, and let $B(\alpha) = \{x \in \mathbb{R}^d \mid ||x||_\infty \leq \alpha\}$ denote the ℓ_∞ ball with radius α. Then the α-inscribed polyhedron of R is the polyhedron

$$R_\alpha := \{x \in R \mid x + B(\alpha) \subset R\}.$$

We are now ready to state our main result. We henceforth fix the notations $P, A, b, W, \mathcal{F}, \mathcal{R}, Q, d$ and n to represent the input to our problem. We denote by Δ and ω the maximum absolute sub-determinant of A, and the largest absolute-value of an entry in W, respectively.

Theorem 7. *$\mathcal{R}$ is $\boldsymbol{\delta}$-regular with respect to the $\boldsymbol{\gamma}$-inscribed polyhedron $Q_{\boldsymbol{\gamma}}$ of Q, where $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ are bounded polynomially in Δ, ω and n.*

Remark 8. *We remark that one can also define a clean notion of a* polyhedral Frobenius number *as follows. Given two matrices $A \in \mathbb{Z}^{m \times n}$ and $W \in \mathbb{Z}^{d \times n}$ let the polyhedral Frobenius number of A and W be*

$$F(A, W) = \min\left\{\max\{\boldsymbol{\gamma}, \boldsymbol{\delta}\} \mid \mathcal{R} \textit{ is } \boldsymbol{\delta}\textit{-regular with respect to } Q_{\boldsymbol{\gamma}}\ \forall b \in \mathbb{Z}^m\right\},$$

where Q, P and $\mathcal{R}$ are defined from A, W and b, as before. We stress that one can define a polyhedral Frobenius number in various alternative ways. With the latter definition, however, Theorem 7 can be restated as follows. The polyhedral Frobenius number $F(A, W)$ is polynomially bounded in Δ, ω and n, and exponentially by d.

The remainder of the paper is organized as follows. In Section 2 we prove Theorem 7. In Section 3 we use Theorem 7 to prove Theorem 2, and mention a number of concrete algorithmic consequences.

2 A proof of Theorem 7

In this section we prove Theorem 7. For that we first introduce some notation and we prove an auxiliary lemma that adapts Theorem 5 to our needs. Then we prove our main theorem.

We start with some notation. Let $B, C \subset \mathbb{R}^d$ and let $D \in \mathbb{R}^{m \times d}$ be a matrix. With $B + C$ we denote the Minkowski sum $\{x \in \mathbb{R}^d \mid x = b + c \text{ with } b \in B \text{ and } c \in C\}$. With DB we denote the set $\{x \in \mathbb{R}^m \mid x = Db \text{ with } b \in B\}$. Further, we denote with $D_{i,\star}$ the i-th row of D and with $D_{\star,i}$ the i-th column. The operator $\lfloor \cdot \rfloor$ maps component-wise every entry to the largest integer smaller than or equal to the corresponding entry. Finally, let $\|D\|_{max} := \max_{i,j} |D_{i,j}|$ denote the maximum absolute value of an entry of D.

Lemma 9. *Let $M \in \mathbb{Z}^{d \times n}$, such that $\text{cone}(M) = \mathbb{R}^d$ and let $\Lambda = M\mathbb{Z}^n$. Let $z \in \Lambda$ and let $\alpha := \|z\|_\infty$. Then z can be expressed as $z = M\lambda$ such that $\lambda \in \mathbb{Z}^n_+$ and $\|\lambda\|_\infty \leq \boldsymbol{\rho}(\alpha, n, \omega)$, where $\boldsymbol{\rho}(\alpha, n, \omega) \in \mathbb{R}_+[x_1, x_2, x_3]$ is a polynomial in α, n and $\omega := \|M\|_{max}$.*

Proof. To start with, we show why we may assume that $\Lambda = \mathbb{Z}^d$. Let B be the Korkin-Zolotarev basis of Λ [15]. Then, a well known property is that the following inequality holds $\|B_{\star,1}\|_2 \cdots \|B_{\star,d}\|_2 \leq ad^d \det(\Lambda)$ (see [16, Theorem 2.3]), where a is a universal constant. It follows that the entries of B are bounded polynomially in $\det \Lambda$. Therefore, there is also a polynomial bound for the entries of its inverse matrix B^{-1}. We can hence transform M, z and Λ by B^{-1} to arrive at a matrix $M' = B^{-1}M$, a vector $z' = B^{-1}z$ and a lattice $\Lambda' = B^{-1}\Lambda$. The entries in M' and z' are polynomially bounded in the entries in M and z, respectively. Furthermore, Λ' becomes the standard lattice, that is $\Lambda' = \mathbb{Z}^d$. We hence assume hereafter that $\Lambda = \mathbb{Z}^d$.

Case 1. Assume that $\{M_{\star,j} \mid j = 1, \ldots, n\} = \{-M_{\star,j} \mid j = 1, \ldots, n\}$, i.e., the negative of every column in M is also a column in M. Let $p = (2^{d-1}\omega^{d-1}, \ldots, 2^0\omega^0)^\mathsf{T}$. Since $\omega = \|M\|_{max}$ and $2^{d-1}\omega^{d-1} \geq \omega \sum_{i=0}^{d-2} 2^i \omega^i$, it holds that $p^\mathsf{T} M_{\star,j} \neq 0$ for all $j = 1, \ldots, n$. Without loss of generality, we assume that n is even, that $p^\mathsf{T} M_{\star,j} > 0$ for all $j = 1, \ldots, n/2$ and that $M_{\star,j} = -M_{\star,n/2+j}$ for $j = 1, \ldots, n/2$. This implies that $\text{cone}(M_{\star,1}, \ldots, M_{\star,n/2})$ is pointed.

By Caratheodory's Theorem (see e.g. [11, Theorem 3.1]), we can express z as a positive combination of at most d linearly independent columns of M, say $z = \sum_{j=1}^d \gamma_{i_j} M_{\star,i_j}$ with $\gamma_{i_j} \in \mathbb{R}_+$. Using Cramer's rule, the Lagrange expansion of determinants and Hadamard's inequality, we can compute the bound

$$\gamma_{i_j} \leq d\alpha\omega^{d-1}(d-1)^{(d-1)/2} =: \boldsymbol{\rho_1}.$$

We set $\gamma_{i_j} = 0$ for $j = d+1, \ldots, n$.

Let $\mathbf{c} = \mathbf{c}((M_{\star,1}, \ldots, M_{\star,n/2}))$, defined as in Theorem 5. Note that $\mathbf{c}$ is polynomially bounded by ω and n. Next, define $\bar{\gamma}_i := \max\{0, \gamma_i - \mathbf{c}\}$ for $i = 1, \ldots, n/2$ and $\bar{\gamma}_i := \min\{\gamma_i, \gamma_{i-n/2} - \mathbf{c}\}$ for $i = n/2+1, \ldots, n$. Let $\bar{z} := M\lfloor\bar{\gamma}\rfloor$. Then $z \in \bar{z} + \mathbf{c}\sum_{j=1}^{n/2} M_{\star,j} + \text{cone}(M_{\star,1}, \ldots, M_{\star,n/2})$. It follows from Theorem 5 that $z-\bar{z}$ can be expressed as a positive integer combination of $M_{\star,1},\ldots,M_{\star,n/2}$. Let $z-\bar{z}=\sum_{j=1}^{n/2}\mu_j M_{\star,j}$, with $\mu_j\in\mathbb{Z}_+$ be such a combination.

Next, we exploit the fact that

$$1\le p^\mathsf{T} M_{\star,j}\le d2^d\omega^d \text{ for all } j\in\{1,\ldots,n/2\}.$$

It holds that $p^\mathsf{T}(z-\bar{z})\le p^\mathsf{T}((\mathbf{c}+1)\sum_{j=1}^{n/2} M_{\star,j})\le(\mathbf{c}+1)n/2d2^d\omega^d$. In particular, this implies that

$$\mu_j\le(\mathbf{c}+1)n/2d2^d\omega^d=:\boldsymbol{\rho_2}.$$

Finally, let $\lambda:=\lfloor\bar{\gamma}\rfloor+\mu$. It holds that $\lambda\in\mathbb{Z}_+^n$, $z=M\lambda$ and $\|\lambda\|_\infty\le\boldsymbol{\rho}(\alpha,n,\omega):=\boldsymbol{\rho_1}+\boldsymbol{\rho_2}$. This completes the proof for Case 1.

Case 2. In the general case $\{M_{\star,j}\mid j=1,\ldots,n\}\ne\{-M_{\star,j}\mid j=1,\ldots,n\}$. Without loss of generality we assume that $-M_{\star,1}\notin\{M_{\star,j}\mid j=1,\ldots,n\}$.

Since $\operatorname{cone}(M)=\mathbb{R}^d$ there exists, by Caratheodory's Theorem, a selection of at most d linearly independent columns of M, such that $-M_{\star,1}=\sum_{j=1}^d\xi_j M_{\star,i_j}$ with $\xi_j\in\mathbb{R}_+$. With $\delta_1:=\det(M_{\star,i_1},\ldots,M_{\star,i_d})$ and $\delta_j=\delta_1\xi_j\in\mathbb{Z}_+$ it follows that

$$-M_{\star,1}=(\delta_1-1)M_{\star,1}+\sum_{j=1}^d\delta_j M_{\star,i_j} \text{ and } \delta_j\le\omega^d d^{d/2}=:\boldsymbol{\rho_3} \tag{3}$$

for all $j=1,\ldots,d$. From this it follows that we can insert $-M_{\star,1}$ to the set $\{M_{\star,j}\mid j=1,\ldots,n\}$ with the slight modification that whenever a multiplier β for column $-M_{\star,1}$ is used in a representation, we replace it by β times its expression for (3). By performing the latter replacement for all n columns independently, we obtain the general bound

$$\|\lambda\|_\infty\le\boldsymbol{\rho}(\alpha,n,\omega):=\boldsymbol{\rho_1}+\boldsymbol{\rho_2}+n\boldsymbol{\rho_2}\boldsymbol{\rho_3}.$$

□

We are now ready to prove Theorem 7

Proof of Theorem 7. In order to distinguish between vectors in fixed dimension d from those in variable dimension n, we denote elements in the n-dimensional space with bold roman letters. We reserve bold greek letters to highlight polynomials. Without loss of generality we assume that P is given in the form $\{\boldsymbol{x}\in\mathbb{R}_+^n\mid A\boldsymbol{x}=b\}$. We can do so by introducing m slack-variables and decomposing any vector into the difference of two nonnegative vectors of the same dimension, i.e. $A\boldsymbol{x}-A\boldsymbol{y}+Iz=b$ with $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}_+^n$ and $z\in\mathbb{R}_+^m$. Note that the dimension only grows linearly and that the maximum absolute subdeterminant remains the same. For each orthant $\mathcal{O}_i$ we define $\mathcal{H}_i$ to be the Hilbert basis of the cone $\{\boldsymbol{x}\in\mathcal{O}_i\mid A\boldsymbol{x}=0\}$. Using Cramer's rule we can express $\{\boldsymbol{x}\in\mathcal{O}_i\mid A\boldsymbol{x}=0\}$ as $\operatorname{cone}(\boldsymbol{g}_i^1,\ldots,\boldsymbol{g}_i^k)$ with $\boldsymbol{g}_i^j\in\mathcal{O}_i\cap\{-\Delta,\ldots,\Delta\}^n$

for $j = 1, \ldots, k$, i.e. the components of the cone generating vectors $\boldsymbol{g}_i^j$ are integral and their absolute values are bounded by Δ. Further, it is well known that $\mathcal{H}_i \subset \{\boldsymbol{x} \in \mathcal{O}_i \cap \mathbb{Z}^n \mid \|\boldsymbol{x}\|_\infty \leq \boldsymbol{\eta_1}\}$ with

$$\boldsymbol{\eta_1} := d\Delta.$$

This implies that for each $\boldsymbol{h} \in \mathcal{H}_i$ it holds that

$$\|W\boldsymbol{h}\|_\infty \leq n\omega d\Delta =: \boldsymbol{\eta_2}.$$

For an introduction to Hilbert bases see, e.g. [18, Section 16.4].

Let $v_1, v_2, \ldots, v_l$ denote the vertices of Q. For each $j = 1, \ldots, l$ there exists a vertex $\boldsymbol{v}_j$ of P such that $W\boldsymbol{v}_j = v_j$. Assuming that $\mathcal{F}$ is non empty we can conclude that for each $j = 1, \ldots, l$ there exists a $\boldsymbol{y}_j \in \mathcal{F}$ such that

$$\|\boldsymbol{v}_j - \boldsymbol{y}_j\|_\infty \leq n\Delta$$

(see [18, Theorem 17.3]). In other words, there exists a feasible integral point $\boldsymbol{y}_j$ close to each vertex $\boldsymbol{v}_j$ for $j = 1, \ldots, l$. Then, with $y_j := W\boldsymbol{y}_j$ it follows that

$$\|v_j - y_j\|_\infty \leq n^2\omega\Delta =: \boldsymbol{\eta_3}.$$

For an illustration of the points in the d-dimensional space see Figure 2.

We will proceed from here as follows. For suitable polynomials $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$, we will consider the pre-image $\boldsymbol{z}$ of an arbitrary point $z \in Q_{\boldsymbol{\gamma}} \cap \mathcal{R}$, construct an affine lattice Λ induced by the Hilbert basis representation of $\boldsymbol{y}_1 - \boldsymbol{z}, \ldots, \boldsymbol{y}_l - \boldsymbol{z}$ containing z, such that $\det(\Lambda) \leq \boldsymbol{\delta}$, and then prove that this lattice intersected with $Q_{\boldsymbol{\gamma}}$ is contained in $\mathcal{R}$. This will then prove our theorem.

Let $z \in Q_{\boldsymbol{\gamma}} \cap \mathcal{R}$ and $\boldsymbol{z} \in \mathcal{F}$ such that $W\boldsymbol{z} = z$. We first exhibit our construction.

For each index $j \in \{1, \ldots, l\}$ we consider the vector $\boldsymbol{y}_j - \boldsymbol{z}$. Let us say that, $\boldsymbol{y}_j - \boldsymbol{z}$ is contained in the orthant $\mathcal{O}_{i_j}$, with $i_j \in \{1, \ldots, 2^n\}$. In view of [19], we can express $\boldsymbol{y}_j - \boldsymbol{z}$ as the positive integer combination of at most $2n - 2$ elements of the Hilbert basis $\mathcal{H}_{i_j}$, i.e.

$$\boldsymbol{y}_j - \boldsymbol{z} = \sum_{k=1}^{2n-2} \lambda_j^k \boldsymbol{h}_j^k,$$

with $\lambda_j^k \in \mathbb{Z}_+$ and $\boldsymbol{h}_j^k \in \mathcal{H}_{i_j}$. Note that all points in $\{\boldsymbol{z} + \sum_{k=1}^{2n-2} \gamma_j^k \boldsymbol{h}_j^k \mid \gamma_j^k \in \mathbb{Z}_+, \text{ and } \gamma_j^k \leq \lambda_j^k\}$ are feasible, i.e. they are a subset of $\mathcal{F}$. This follows from the fact that $\boldsymbol{h}_j^k \in \{\boldsymbol{x} \in \mathcal{O}_{i_j} \mid A\boldsymbol{x} = 0\}$ for every $k = 1, \ldots, 2n - 2$.

Let

$$\boldsymbol{\rho} := \boldsymbol{\rho}((2n-2)l\boldsymbol{\eta_1}, (2n-2)l, \boldsymbol{\eta_2})$$

be the polynomial defined in Lemma 9. Letting $\boldsymbol{\eta_4} := (2n-2)l(\mathbf{p}+1)\boldsymbol{\eta_1}$, we define

$$\bar{\lambda}_j^k := \max\{0, \lambda_j^k - \boldsymbol{\eta_4}\}$$

and

$$\bar{\boldsymbol{y}}_j := \boldsymbol{z} + \sum_{k=1}^{2n-2} \bar{\lambda}_j^k \boldsymbol{h}_j^k. \tag{4}$$

Notice that $\bar{y}_j := W\bar{\boldsymbol{y}}_j$ remains close to its corresponding vertex v_j. That is

$$\|v_j - \bar{y}_j\|_\infty \leq \boldsymbol{\eta_3} + (2n-2)\boldsymbol{\eta_2}\boldsymbol{\eta_4} =: \boldsymbol{\eta_5}.$$

Choosing $\boldsymbol{\gamma} \geq \boldsymbol{\eta_5}$ we ensure that z is sufficiently far from each vertex v_j so that at least one λ_j^k must be greater or equal than $\boldsymbol{\eta_4}$ for each j.

For simplicity, we assume that $\bar{\lambda}_j^k > 0$ for all j and k. This can be assumed without loss of generality, as if $\lambda_j^k = 0$ we can simply modify $\boldsymbol{y}_j$ and consider a representation of it with one Hilbert basis element less. Let $h_j^k := W\boldsymbol{h}_j^k$ for every $j = 1, \dots, 2n-2$ and $k = 1, \dots, l$. We define the affine lattice

$$\Lambda = \{x \in \mathbb{Z}^d \mid x = z + \sum_{j=1}^{l} \sum_{k=1}^{2n-2} \gamma_j^k h_j^k,\ \gamma_j^k \in \mathbb{Z},\ \forall j,k\}.$$

We can bound the determinant of Λ by any determinant of any sub-lattice induced by d linearly independent h_j^i-s. Hence, by Hadamard's inequality and since $\|h_j^k\|_\infty \leq \boldsymbol{\eta_2}$, it holds that $\det(\Lambda) \leq \boldsymbol{\delta} := \boldsymbol{\eta}_2^d d^{d/2}$. Next, we define the matrix

$$M := [h_1^1, \dots, h_1^{(2n-2)}, \dots, h_l^1, \dots, h_l^{(2n-2)}].$$

It holds that $\|M\|_{max} \leq \boldsymbol{\eta_2}$. In order to apply Lemma 9 let us first verify Claim 1.

Claim 1. $\text{cone}(M) = \mathbb{R}^d$.

Proof of Claim 1. Assume that $\text{cone}(M) \neq \mathbb{R}^d$. Then there exists a $u \in \mathbb{R}^d$ with $\|u\|_2 = 1$ defining a half-space $\{x \in \mathbb{R}^d \mid u^\intercal x \leq 0\}$ such that $\text{cone}(M) \subset \{x \in \mathbb{R}^d \mid u^\intercal x \leq 0\}$. Since $z \in Q_{\boldsymbol{\gamma}}$ it holds that $z + B(\boldsymbol{\gamma}) \subset Q$. This implies that there exists a vertex v_i such that $u^\intercal v_i - u^\intercal z \geq \boldsymbol{\gamma}$. On the one hand, it holds that

$$\boldsymbol{\gamma} \leq u^\intercal(v_i - y_i + y_i - z) \leq \boldsymbol{\eta_3} + \sum_{k=1}^{2n-2} \lambda_i^k u^\intercal h_i^k.$$

On the other hand, for each $k = 1, \dots, 2n-2$ it holds that $u^\intercal h_i^k \leq \|h_i^k\|_\infty \leq \boldsymbol{\eta_2}$. It follows that for some $j \in \{1, \dots, 2n-2\}$ we have that $u^\intercal h_i^j > 0$ and $\lambda_i^j \geq \boldsymbol{\eta_4}$. By construction, this implies that h_i^j is a column of M, contradicting that $\text{cone}(M) \subset \{x \in \mathbb{R}^d \mid u^\intercal x \leq 0\}$. □

It remains to show that $\mathcal{R}$ is $\boldsymbol{\delta}$-regular with respect to $Q_{\boldsymbol{\gamma}}$. We split the proof into Claim 2 and 3. In Claim 2 we show that for every $j = 1, \dots, l$, all lattice points sufficiently close to $\bar{y}_j$ have a feasible pre-image.

Claim 2. For every $\gamma \in \mathbb{Z}_+^{(2n-2)l}$ with $\|\gamma\|_\infty \leq \boldsymbol{\rho} + 1$ it holds that $\bar{y}_j + M\gamma$ has a feasible pre-image, i.e. it is not a hole.

Proof of Claim 2. We prove the claim by showing that $\bar{\boldsymbol{y}}_j + (2n-2)l(\boldsymbol{\rho}+1)\boldsymbol{h}_i^k \in P$ for every i, j and k. This will then imply the slightly stronger result, that

$$\bar{y}_j + M\gamma$$

is feasible for any $\gamma \in \mathbb{Z}_+^{(2n-2)l}$ with $\|\gamma\|_1 \leq (2n-2)l(\boldsymbol{\rho}+1)$.

In order do derive a contradiction, let us assume that the latter does not hold for $j=1$ and $i=2$, i.e. that $\bar{\boldsymbol{y}}_1 + (2n-2)l(\boldsymbol{\rho}+1)\boldsymbol{h}_2^1 \notin P$. The only constraints defining P that can be violated by this vector are the non-negativity constraints, thus some component of this vector must be strictly negative. Let us assume that the first component is negative. We have an upper and a lower bound for this component, namely

$$-(2n-2)l(\boldsymbol{\rho}+1)\boldsymbol{\eta_1} \leq (\boldsymbol{y}_1 + (2n-2)l(\boldsymbol{\rho}+1)\boldsymbol{h}_2^1)_1 < 0.$$

All the vectors $\boldsymbol{h}_1^k$, $k = 1, \dots, 2n-2$, lie in the same orthant, therefore for each i all the entries $(\boldsymbol{h}_1^k)_i$, $k = 1, \dots, 2n-2$, must either be all non-positive or all non-negative. Since $\boldsymbol{z} + (2n-2)l(\boldsymbol{\rho}+1)\boldsymbol{h}_2^1$ is feasible, it must hold that $(-\boldsymbol{h}_1^k)_1 \geq 0$ for all $k = 1, \dots, 2n-2$. In particular, there must be at least one index $k \in \{1, \dots, 2n-2\}$ such that $(-\boldsymbol{h}_1^k)_1 > 1$. Hence,

$$0 \leq \left(\boldsymbol{y}_1 + (2n-2)l(\boldsymbol{\rho}+1)\boldsymbol{h}_2^1 - \boldsymbol{\eta_4}\sum_{k=1}^{2n-2}\boldsymbol{h}_1^k\right)_1 = \left(\bar{\boldsymbol{y}}_1 + (2n-2)l(\boldsymbol{\rho}+1)\boldsymbol{h}_2^1\right)_1 < 0.$$

We obtained a contradiction. □

We now use Claim 2 to show that all points in $\Lambda \cap Q_{\boldsymbol{\gamma}}$ have pre-images, i.e. they are not holes.

Claim 3. $\Lambda \cap Q_{\boldsymbol{\gamma}} \subset \mathcal{R}$.

Proof of Claim 3. Let $\bar{z} \in \Lambda \cap Q_{\boldsymbol{\gamma}}$. We prove that there exists a $\bar{\boldsymbol{z}} \in P \cap \mathbb{Z}^n$ such that $\bar{z} = W\bar{\boldsymbol{z}}$. Note that $Q_{\boldsymbol{\gamma}} \subset \operatorname{conv}(\bar{y}_1, \dots, \bar{y}_l)$. By Caratheodory's theorem there exist $i_1, \dots, i_d \in \{1, \dots, l\}$, such that $\bar{z} \in \operatorname{conv}(z, \bar{y}_{i_1}, \dots, \bar{y}_{i_d})$. Without loss of generality we may assume that $i_1 = 1, \dots, i_d = d$. Let $\alpha \in [0,1]$ and $\alpha_j \in \mathbb{R}_+$ for $j = 1, \dots, d$, such that $\bar{z} = \alpha z + (1-\alpha)\sum_{j=1}^d \alpha_j \bar{y}_j$ and $\sum_{j=1}^d \alpha_j = 1$. Hence, using (4), $\bar{z}$ is the image under W of a not necessarily integral point

$$\boldsymbol{z} + (1-\alpha)\sum_{j=1}^{d}\alpha_j \sum_{k=1}^{2n-2}\bar{\lambda}_j^k \boldsymbol{h}_j^k,$$

which is included in P. We can approximate this point by

$$\hat{\boldsymbol{z}} = \boldsymbol{z} + \sum_{j=1}^{d}\sum_{k=1}^{2n-2}\lceil(1-\alpha)\alpha_j\bar{\lambda}_j^k\rceil \boldsymbol{h}_j^k.$$

Clearly, $\hat{z} := W\hat{\boldsymbol{z}} \in \Lambda$. Let $\boldsymbol{l} := \sum_{j=1}^d \sum_{k=1}^{2n-2}(\lceil(1-\alpha)\alpha_j\bar{\lambda}_j^k\rceil - (1-\alpha)\alpha_j\lambda_j^k)\boldsymbol{h}_j^k$. Since $\boldsymbol{z} + \boldsymbol{l} \in P$ and $\boldsymbol{y}_j + \boldsymbol{l} \in P$ holds for all $j = 1, \dots, d$ (see Claim 2), $\hat{\boldsymbol{z}}$

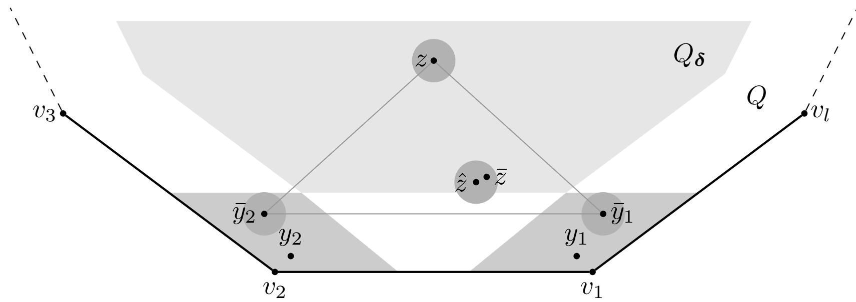

Figure 2: An illustration of the notation in the proof of Theorem 7.

must be feasible. From Claim 2 it follows again that $\hat{z} + M\gamma$ is feasible for any $\gamma \in \mathbb{Z}_+^{2nl-2l}$ with $\|\gamma\|_\infty \leq \boldsymbol{\rho}$. It holds that $\|\bar{z} - \hat{z}\|_\infty \leq (2n-2)l\boldsymbol{\eta_1}$. Finally, we can apply Lemma 9 to guarantee the existence of $\bar{\boldsymbol{z}} \in \mathcal{F}$ such that $\bar{z} = W\bar{\boldsymbol{z}}$. □

This completes the proof of the theorem. □

3 Applications to non-linear integer optimization

We describe next a general algorithmic framework that allows us to apply Theorem 7 to solve variable-dimension non-linear integer optimization problems. More precisely, we show a general purpose algorithm that solves Problem (1) with a number of oracle calls that is polynomial in the input size, Δ, n and ω, thus proving Theorem 2. For brevity, we will henceforth say "in polynomial time" to imply a running time of the latter type. We recall that the oracles available to our algorithm are an optimization oracle and a fiber oracle (see Definition 1). We stress that the dependence on d can be exponential. We assume in this section that d is an arbitrary fixed constant. We later mention a number of concrete examples of problem classes, for which, using polynomial-time implementations of the oracles, our algorithm runs in polynomial time in the encoding length of the input.

To simplify notation, we will henceforth restrict our attention to minimization problems. We stress that Theorem 2 works also for maximization problems. Our algorithm works with an inequality description of the polyhedra Q and $Q_{\boldsymbol{\gamma}}$. Since the input only provides implicit representations of these polyhedra, we need the following lemma, which also states a useful connection between the two descriptions.

Lemma 10. *One can compute in polynomial time a matrix $F \in \mathbb{Z}^{q\times d}$ and vectors $g, g' \in \mathbb{Z}^q$ such that $Q = \{x \in \mathbb{R}^d \mid Fx \leq g\}$ and $Q_{\boldsymbol{\gamma}} = \{x \in \mathbb{R}^d \mid Fx \leq$*

$g'\}$*, with* $\|F\|_{max} \leq (n\omega\Delta)^{d-1}(d-1)^{(d-1)/2}$ *and*

$$|g_i - g'_i| \leq \boldsymbol{\gamma}\|F_{i,\star}\|_\infty$$

for every $i \in [q]$.

Proof. We start with some notation. Let v_1 and v_2 denote two adjacent vertices of Q. Together they define the edge $\text{conv}(v_1, v_2)$ of Q. In the following we call a vector e an *edge-direction* of an edge $\text{conv}(v_1, v_2)$, if $e \in \text{lin}(v_2 - v_1)$.

To prove the lemma, we exploit that each edge-direction of Q is the image (under the linear mapping W) of an edge-direction of P. An edge-direction E of P, which corresponds to an edge-direction of Q, can be expressed as the intersection of $n-1$ linearly independent facets. Let us assume without loss of generality that these are $A_{1,\star}, \ldots, A_{n-1,\star}$. Applying Cramer's rule we know that there exists a non-trivial solution $E \in \mathbb{Z}^n$ such that $A_{i,\star}E = 0$ for all $i \in \{1, \ldots, n-1\}$ and $\|E\|_\infty \leq \Delta$. Let $e := WE$. It follows that, $\|e\|_\infty \leq n\omega\Delta$.

A facet of Q is defined by $d-1$ linear independent edge-directions, say $e_1, \ldots, e_{d-1}$. Then, a facet defining vector $F_{i,\star}$ is defined by a non-trivial solution to $e_i^\mathsf{T} x = 0$. Using Cramer's rule and Hadamard's inequality we can choose $F_{i,\star} \in \mathbb{Z}^d$ such that $\|F_{i,\star}^\mathsf{T}\|_\infty \leq (n\omega\Delta)^{d-1}(d-1)^{(d-1)/2}$. It remains to note that there is only a polynomial number of possible $F_{i,\star}$. Hence, one can compute F and g by brute force with linear programming [18].

We can now find an inequality description of $Q_{\boldsymbol{\gamma}}$ as follows. First, by normalizing the inequalities defining Q, i.e., by setting $\bar{F}_{i,\star} = \frac{1}{\|F_{i,\star}\|_\infty}F_{i,\star}$ and $\bar{g}_i = \frac{1}{\|F_{i,\star}\|_\infty}g_i$, one easily verifies that

$$Q_{\boldsymbol{\gamma}} = \{x \in \mathbb{R}^d \mid \bar{F}_{i,\star}^\mathsf{T} x \leq \bar{g}_i - \boldsymbol{\gamma} \quad \forall i \in [q]\}.$$

A description with integral coefficients is hence given by

$$Q_{\boldsymbol{\gamma}} = \{x \in \mathbb{R}^d \mid F_{i,\star}^\mathsf{T} x \leq g_i - \boldsymbol{\gamma}\|F_{i,\star}\|_\infty \quad \forall i \in [q]\},$$

so we can set $g'_i = g_i - \boldsymbol{\gamma}\|F_{i,\star}\|_\infty$ for all $i \in [q]$. The bound $|g_i - g'_i| \leq \boldsymbol{\gamma}\|F_{i,\star}\|_\infty$ immediately follows. ☐

As was discussed in the introduction, our algorithmic approach relies on a decomposition of the problem into "sufficiently regular" sub-problems. Each sub-problem corresponds to a projected feasible set $R \cap \Lambda \subset \mathbb{Z}^d$ containing no holes, where R is a polyhedron and Λ is a lattice. Then, the optimization oracle is invoked to obtain a point $y^* \in R \cap \Lambda$ attaining

$$\min\{f(y) \mid y \in R \cap \Lambda\},$$

and a point $x^* \in \mathcal{F}$ is computed with $Wx^* = y^*$ using the fiber oracle. The best solution across all sub-problems is then an optimal solution.

We distinguish between two types of sub-problems. The first type is concerned with the polyhedron $Q_{\boldsymbol{\gamma}}$, i.e., such sub-problems optimize over the restricted feasible region

$$\mathcal{F}' := \{x \in \mathcal{F} \mid Wx \in Q_{\boldsymbol{\gamma}}\}.$$

In the following lemma we prove, using Theorem 7, that the optimal point in this region can be found efficiently.

Lemma 11. *The problem*

$$\min \{f(Wx) \mid x \in \mathcal{F}'\}$$

can be solved with a polynomial number of calls to the optimization and fiber oracles.

Proof. As guaranteed by Theorem 7, for every point $x \in \mathcal{F}'$ there is a lattice Λ_x with determinant at most $\boldsymbol{\delta}$ such that $x \in \Lambda_x$ and $\Lambda_x \cap Q_{\boldsymbol{\gamma}} \subset \mathcal{R}$, i.e., $\Lambda_x \cap Q_{\boldsymbol{\gamma}}$ contains no holes. Consider an optimal solution y^* to the problem

$$\min\{f(y) \mid y \in \Lambda_x \cap Q_{\boldsymbol{\gamma}}\},$$

obtainable by a single oracle call to the optimization oracle. Since $\Lambda_x \cap Q_{\boldsymbol{\gamma}}$ contains no holes, one can obtain, using a call to the fiber oracle, a pre-image $x^* \in \mathcal{F}'$ of y^*. Furthermore, due to $x \in \Lambda_x \cap Q_{\boldsymbol{\gamma}}$ we also know that $f(x^*) \leq f(x)$. Consequently, to minimize over $\mathcal{F}'$ it suffices to consider the problem

$$\min\{f(y) \mid y \in \Lambda \cap Q_{\boldsymbol{\gamma}}\},$$

for every affine lattice Λ with determinant bounded by $\boldsymbol{\delta}$. Next, we bound the number of such lattices.

An affine lattice can be represented by a basis $B \subset \mathbb{Z}^{d\times d}$ and a translation vector $t \in \{B\lambda \mid \lambda \in [0,1)^d\} \cap \mathbb{Z}^d$ as

$$\Lambda = \{t + v \mid \exists z \in \mathbb{Z}^d \;\; v = Bz\}.$$

We can assume that B comprises the columns of a matrix in Hermite Normal Form. The bound on the determinant of the lattice now translates to a bound on the maximum absolute value of an entry in B. We can thus roughly estimate the number of affine lattices by $\boldsymbol{\delta}^{d^2+d}d^d$.

It follows that, by considering every bounded-determinant lattice, as described before, one can obtain the best solution $x \in \mathcal{F}'$ with at most $2\boldsymbol{\delta}^{d^2+d}d^d$ oracle calls. □

To treat the region $\mathcal{F} \setminus \mathcal{F}'$ we use a recursive decomposition into lower-dimensional problems. To control the number of such problems we use the fact that all points $y = Wx$ for points $x \in \mathcal{F} \setminus \mathcal{F}'$ fall close to the boundary of Q. This fact is used in the following lemma to prove a bound on the number of hyperplanes needed to cover all integer points in $Q \setminus Q_{\boldsymbol{\gamma}}$.

Lemma 12. *There is a polynomial time procedure that computes a set $\mathcal{H}$ of hyperplanes parallel to the facets of Q, with the property that all integer points in $Q \setminus Q_{\boldsymbol{\gamma}}$ lie on at least one hyperplane in $\mathcal{H}$, i.e.,*

$$(Q \setminus Q_{\boldsymbol{\gamma}}) \cap \mathbb{Z}^d \subset \bigcup_{H\in\mathcal{H}} H$$

In particular, $\mathcal{H}$ has polynomial size.

Proof. Lemma 10 asserts that Q and $Q_{\boldsymbol{\gamma}}$ admit inequality descriptions $Q = \{x \in \mathbb{R}^d \mid Fx \leq g\}$ and

$$Q_{\boldsymbol{\gamma}} = \{x \in \mathbb{R}^d \mid F_{i,\star}^{\mathsf{T}} x \leq g_i - \boldsymbol{\gamma}\|F_{i,\star}\|_\infty \ \ \forall i \in [q]\},$$

with $\|F_{i,\star}\|$ polynomially bounded for all $i \in [q]$. We can now use this description to cover all integer points in $Q \setminus Q_{\boldsymbol{\gamma}}$ with a polynomial number of hyperplanes parallel to the facets of Q. More precisely, for $i \in [q]$ let $\mathcal{H}_i$ denote the set of hyperplanes of the form

$$H_i(s) = \{x \in \mathbb{R}^d \mid F_{i,\star}^{\mathsf{T}} x = s\},$$

where $s \in \{g_i - \boldsymbol{\gamma}\|F_{i,\star}\|_\infty, g_i - \boldsymbol{\gamma}\|F_{i,\star}\|_\infty + 1, \cdots, g_i\}$ ranges over all integer right-hand sides between $g_i - \boldsymbol{\gamma}\|F_{i,\star}\|_\infty$ and g_i. Note that, by Lemma 10, the number of such hyperplanes is indeed polynomially bounded. We can now take the union over all facets of Q of the sets $\mathcal{H}_i$, i.e.

$$\mathcal{H} = \bigcup_{i \in [q]} \mathcal{H}_i$$

to arrive at the desired set of hyperplanes. Since the number of facets of Q is polynomially bounded, the lemma is proved.

□

We now have almost all ingredients for the proof of Theorem 2. The following remark states that the constraint matrix of sub-problems arising by restricting the feasible set to the pre-image of an arbitrary face of Q has a determinant that is polynomially bounded.

Remark 13. *Let $I \subset \mathbb{R}^d$ denote an i-face of Q. Let $F_{i_1,\star}, \ldots, F_{i_{d-i},\star} \in \mathbb{Z}^d$ be the facets defining the face Q, i.e. $I = \{x \in Q \mid F_{i_j,\star}^{\mathsf{T}} x = g_{i_j} \ \ \forall j \in [d-i]\}$. Then $W^{-1}I$, the pre-image of I under W, can be expressed as $\{x \in \mathbb{R}^n \mid \bar{A}x \leq \bar{b}\}$ with*

$$\bar{A} := [A^{\mathsf{T}}, (F_{i_1,\star}W)^{\mathsf{T}}, -(F_{i_1,\star}W)^{\mathsf{T}}, \ldots, (F_{i_{d-i},\star}W)^{\mathsf{T}}, -(F_{i_{d-i},\star}W)^{\mathsf{T}}]^{\mathsf{T}}$$

and

$$\bar{b} := [b^{\mathsf{T}}, g_{i_1}, -g_{i_1}, \ldots, g_{i_{d-i}}, -g_{i_{d-i}}]^{\mathsf{T}}.$$

In particular, note that the maximum absolute sub-determinant of $\bar{A}$ is polynomially bounded.

We are now ready to prove Theorem 2.

Proof of Theorem 2. The algorithm starts by computing the inequality descriptions of Q and $Q_{\boldsymbol{\gamma}}$, as in Lemma 10. Then, the algorithm proceeds by solving the problem

$$\min\{f(Wx) \mid x \in \mathcal{F}'\}$$

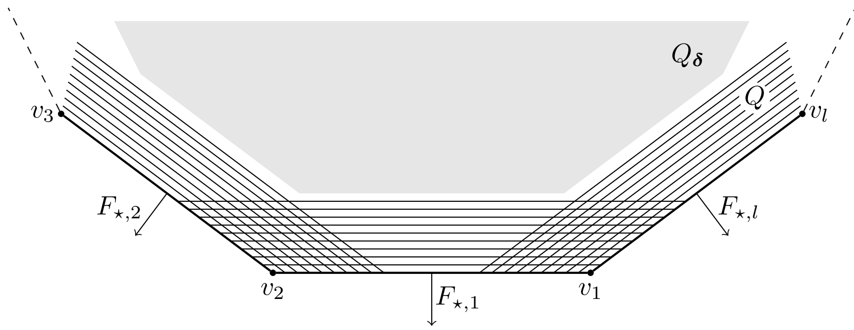

Figure 3: An illustration of the algorithm.

by invoking the procedure in Lemma 11. To treat the region $\mathcal{F} \setminus \mathcal{F}'$, the algorithm obtains first the polynomially-bounded set of hyperplanes $\mathcal{H}$, using the procedure in Lemma 12. For every hyperplane $H \in \mathcal{H}$, the algorithm recursively solves the $(d-1)$-dimensional problem

$$\min \{ f(Wx) \mid x \in \mathcal{F}(H) \} ,$$

where $\mathcal{F}(H) := \{x \in \mathcal{F} \mid Wx \in H \cap Q\}$. Each such sub-problem admits the same form as the original one. Additionally, Remark 13 implies that the matrix corresponding to the inequality description of $P' := \{x \in \mathbb{R}^d \mid Wx \in H \cap Q\}$ has determinant that is polynomially bounded, as well. See Figure 3 for an illustration of the algorithm.

Finally, since d is fixed, so is the depth of the recursion, implying that the algorithm performs, in total, a polynomial number of oracle calls, and additional polynomial work. $\square$

Theorem 2 achieves our main algorithmic goal, namely a general-purpose efficient reduction from non-linear integer programming to linear integer programming. The linear integer programs that arise correspond to the fiber problem, implying that their feasible set is defined from the matrices given in the input data. This property is desirable, since our algorithm does not require solving linear integer programs with a feasible set, whose structure dramatically differs from that of the original non-linear problem. Consequently, our reduction makes it possible to solve a large class of non-linear integer problems using well-known techniques for linear integer programming, such as cutting planes methods etc.

We conclude the paper by mentioning some concrete class of problems solved by our algorithm. Unless stated otherwise, no polynomial time algorithms were known for these problems. To arrive at the desired polynomial algorithms we need to present polynomial implementations of the optimization and fiber oracles. Let us first list a number classes of non-linear functions for which the optimization oracle can be implemented in polynomial time. We stress that the latter results hold in fixed dimension, i.e., whenever d is an arbitrary, but fixed constant. In all cases the feasible set comprises an arbitrary intersection of a polyhedron and an affine lattice, whose descriptions are provided in the input, and the functions are presented with evaluation oracles. We stress that any combination of implementable oracles from the lists below lead to a class of optimization problems solvable by our algorithm.

- **Minimization of convex functions.** Grötschel, Lovász and Schrijer [10] presented an algorithm for the minimization of a convex function presented by evaluation oracles.

- **Minimization of bounded degree polynomials.** Del Pia and Weismantel [7] presented an algorithm for minimizing arbitrary degree-two polynomials with integer coefficients in the plane. This result was recently extended by Del Pia, Hildebrand, Weismantel and Zemmer [6] to cubic polynomials in two variables, in the case of a bounded polyhedron. With the same restriction on the feasible set, the authors also present a polynomial algorithm for minimizing a homogeneous polynomial with two variables and an arbitrary fixed degree.

- **Approximate maximization of non-negative polynomials.** De Loera, Hemmecke, Köppe and Weismantel [4] showed that a polynomial in fixed dimension can be approximately maximized in polynomial time over the integer points in a polyherdon, provided that the polynomial is non-negative over the polyhedron. Concretely, the authors show a fully polynomial-time approximation scheme (FPTAS) for the problem.

We stress that the latter list gives a few prominent examples of classes for which the optimization oracle can be implemented efficiently, but it is far from being a complete list. We remark that in order to obtain an approximate solution to Problem (1) it suffices to employ an approximate implementation of the optimization oracle.

We turn to implementations of the fiber oracle. Recall that the fiber oracle is required to provide a point in $\{x \in \mathbb{Z}^n : Ax \leq b,\ Wx = y\}$ for an arbitrary $y \in \mathbb{Z}^d$, if one exists, or correctly report that the latter set is empty.

- **A Constant number of constraints.** Eisenbrand, Vempala and Weismantel [8] recently showed that an integer program with a fixed number of rows can be solved in time polynomial in the dimension and the maximum sub-determinant of the constraint matrix, and *independent* of the right-hand side. This result implies that when $\binom{W}{A}$ has a constant number of rows, and the entries in this matrix are polynomially bounded in the input length, the fiber oracle can be implemented in polynomial time.

- **N-fold systems.** It is well-known that if A is an N-fold matrix then the matrix $\binom{W}{A}$ can be transformed to an equivalent N-fold matrix, provided that all entries in W form a set $K \subset \mathbb{Z}$ of fixed size. As was shown by De Loera, Hemmecke, Köppe and Weismantel [5], integer programs with an N-fold constraint matrix admit polynomial-time algorithms.

We note that there are several other interesting classes of matrices that admit polynomial algorithms. One obvious example is when $\binom{W}{A}$ is totally unimodular. In such cases, however, one has $\mathcal{R} = Q \cap \mathbb{Z}^d$, so the Problem (1) can in these cases be solved with two oracle calls (see Remark 3).

As a final note let us state one concrete new implication of Theorem 2.

Corollary 14. *Let m, d be some fixed positive integers, and let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function presented by an evaluation oracle. There is a polynomial algorithm that accepts the unary representation of two matrices $A \in \mathbb{Z}^{m \times n}$, $W \in \mathbb{Z}^{d \times n}$ and the binary representation of a vector $b \in \mathbb{Z}^m$ and solves the problem*

$$\min \{ f(Wx) \mid Ax \leq b,\ x \in \mathbb{Z}^n \} .$$

Proof. Since the number of rows in A is fixed, the maximum sub-determinant of A is bounded by a polynomial in the maximum entry in A which, due to the unary encoding of A, is polynomially bounded by the input length. Furthermore, the maximum entry in W is polynomially bounded by the input length.

The result now immediately follows from Theorem 2 and the aforementioned results of Grötschel, Lovász and Schrijer [10] and Eisenbrand, Vempala and Weismantel [8]. □

References

[1] I. Aliev and M. Henk. Feasibility of integer knapsacks. *SIAM J. Optim.*, 20(6):2978–2993, 2010.

[2] A. Brauer. On a problem of partitions. *American Journal of Mathematics*, 64(1):299–312, 1942.

[3] A. Brauer and J. E. Shockley. On a problem of Frobenius. *J. reine angew. Math*, 211:215–220, 1962.

[4] J. A. De Loera, R. Hemmecke, M. Köppe, and R. Weismantel. Integer polynomial optimization in fixed dimension. *Mathematics of Operations Research*, 31:147–153, 2006.

[5] J. A. De Loera, R. Hemmecke, S. Onn, and R. Weismantel. N-fold integer programming. *Discrete Optimization*, 5(2):231–241, 2008.

[6] A. Del Pia, R. Hildebrand, R. Weismantel, and K. Zemmer. Minimizing cubic and homogeneous polynomials over integers in the plane. Manuscript, 2014.

[7] A. Del Pia and R. Weismantel. Integer quadratic programming in the plane. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 840–846, 2014.

[8] F. Eisenbrand, S. Vempala, and R. Weismantel. Revisiting dynamic programming for integer optimization. Private communication, 2014.

[9] P. Erdös and R. L. Graham. On a linear Diophantine problem of Frobenius. *Acta Arithmetica*, 21:399–408, 1972.

[10] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, volume 2 of *Algorithms and Combinatorics: Study and Research Texts*. Springer-Verlag, Berlin, 1988.

[11] P.M. Gruber. *Convex and Discrete Geometry*, volume 336 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2007.

[12] D. S. Hochbaum and J. G. Shanthikumar. Convex separable optimization is not much harder than linear optimization. *J. ACM*, 37(4):843–862, 1990.

[13] R. Kannan. Solution of the Frobenius problem and its generalization. Manuscript, 1989.

[14] R. Kannan. Lattice translates of a polytope and the Frobenius problem. *Combinatorica*, 12(2):161–177, 1992.

[15] A. Korkin and G. Zolotarev. Sur les formes quadratiques. *Math. Ann.*, 6(3):366–389, 1873.

[16] J. C. Lagarias, H. W. Lenstra, Jr., and C.-P. Schnorr. Korkin-Zolotarev bases and successive minima of a lattice and its reciprocal lattice. *Combinatorica*, 10(4):333–348, 1990.

[17] J.L. Ramírez-Alfonsín. Complexity of the Frobenius problem. *Combinatorica*, 16(1):143–147, 1996.

[18] A. Schrijver. *Theory of linear and integer programming*. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons Ltd., Chichester, 1986. A Wiley-Interscience Publication.

[19] A. Sebö. Hilbert bases, Caratheodory's theorem and combinatorial optimization. In *Proceedings of the 1st Integer Programming and Combinatorial Optimization Conference*, pages 431–455. University of Waterloo Press, 1990.